Exploring events along the bacterial Twin arginine Transport pathway

Shanmugham, A.

2011

Outlook

Using Artificial substrates to study protein transport processes.

Green fluorescent protein (GFP) has been of undeniable benefit to thousands of cell biologists. The major advantages of using GFP are their ability to readily and firmly fold upon expression in cells and then display intrinsic fluorescence. Fluorescence is endowed by three amino acids Ser_{65}-Tyr_{66}-Gly_{67} in GFP that form the fluorophore following a complex maturation process. The introduction of GFP as epitope tags onto target protein has long been a valuable way of assessing the dynamics and subcellular localization of the target protein. In this thesis (chapter 1, 3 and 5), we distinctively employ GFP based artificial transport substrates in both *in vivo* and *in vitro* studies involving protein transport processes across bacterial membranes. The tools illustrated here have wider applications beyond the scope of this thesis.

The production of transport-competent substrates in chemical amounts were made possible by employing a self-cleaving N-terminal intein tag that allows chromatographic purification and even more importantly cap and protect the signal during expression and purification. We use this strategy to produce GFP fused to different (both sec and tat) signal peptides in the N-terminus. We have demonstrated the potential of these artificial transport substrates in understanding signal sequence interaction with various Tat transport pathway constituents: phospholipids (in Chapter 2) and chaperones (in Chapter 3). The next logical step employing these artificial transport substrates is to unravel the particular role of the Tat signal peptide in mediating binding of the substrates to the Tat complex components: mainly TatA, TatB, TatC. Potentially, much insight could be gained from *in vitro* experiments employing purified Tat components in interaction studies under controlled conditions. For such purpose, Tat translocase from a hyperthermophile such as *Aquifex aeolicus* can be used as model system in the place of relatively less stable *Escherichia coli* Tat translocase. In chapter 4, we have already shown *in vivo* the ability of *tat* genes from the *A. aeolicus*, to compensate for the absence of the cognate *E. coli* tat genes and therefore *A. aeolicus* Tat complex is expected to have transport related properties similar to that of *E.coli* Tat complex. Stable transport complexes also have a good chance to be purified at suitable levels required for *in vitro* purposes.

In chapter 3, using GFP based artificial substrates we confirmed that, for REMPs, TorD and DmsD, a substrate's signal sequence alone suffices for tight and specific binding to the correct REMP. The X-ray structures of TorD from *Shewanella massilia* (Tranier et al, 2003) and DmsD homologues from *E. coli* (Stevens et al, 2009), *S. typhimurium* LT2 (Qiu et al, 2008) and *Archaeoglobus fulgidus* (Kirillova et al, 2007) have been reported. All these REMPs show highly similar all-α-helical structures. It is thus difficult to predict where in the signal sequence

or REMP the specificity exactly resides. The signal sequences of TorA and DmsA are rather similar as well (Chapter 2). They share the same overall tripartite structure of a signal sequence, consisting of a positively charged N-region, a hydrophobic H-region and a polar C-region. Furthermore, the H-regions are similar in length and in average hydrophobicity. The C-regions are also of similar length, while both contain a proline and at least one arginine (ssTorA has two, ssDmsA one). The main difference between the two signal sequences lies in the N-regions. The N-region of ssDmsA is five residues longer than the one of TorA. Furthermore, the N-region of ssDmsA is significantly more hydrophobic. Overall, it seems important to find out in which region of the signal peptide is the specificity for REMP binding encoded. In this case, the tool box of artificial transport substrates can be extended by generating substrates with chimeric signal sequences based on ssTorA-GFP and ssDmsA-GFP, in which either the N- or the H- or the C-regions of one signal sequence were replaced by the corresponding sequence of the other signal. Such preproteins can be screened qualitatively for interaction with DmsD and TorD using SPR.

Substrates with mutations in the signal peptide region are also shown to be useful tools to test the relevance of specific features to their function. For instance, to test whether the double arginine motif in the Tat signal peptide has any role in the binding of REMPs, we mutated either one or both of the two arginine residues (RR) in the TorA signal sequence (see Chapter 3). The result from the mutated substrates clearly showed that the conserved arginine pair in the signal peptides of Tat substrates does not play an important role in binding of the peptides to the cognate REMPs. Yet another interesting scenario is a mutation of a particular proline residue in the H- region of the signal peptide. The proline at position 29 in the ssTorA is absent in ssDmsA and this seems to be one most striking difference between both signal. Provided that indeed the H-region of a signal peptide adopts an α-helix upon binding to the hydrophobic binding pocket on the surface of a REMP as recently studied by Chan et al, 2008 and Zakian et al, 2010, a proline residue in the H-region would result in a kinked α-helix. This kink might be what causes specificity in the interaction between a signal peptide and a REMP. In order to check this out, the proline at position 29 could be replaced with an alanine, resulting in a pre-protein ssTorA(P29A)-GFP and screened for REMP interaction.

The application of Intein that was initially meant to be a purification tag took an unexpected twist in our study after we observed that by fusing a cleavable intein N-terminally to signal GFPs we actually intervene putative co-translational sorting process (evidences in Chapter 5). We therefore employed the intein-tagged GFP/cytochrome *c* based substrates in *in vivo* experiments in order to address the questions “where, when and how Sec substrates are sorted from Tat substrates”. As a result, we show that the Tat system rescues the transport of folded Sec substrates, and we present evidence that the sorting between Sec and Tat substrates can be performed at a late post-translational stage, probably upon interaction with the SecYEG pore. These intein tagged- GFP based substrates can also be applied in other related studies and

one area that we are particularly tempted to use this is towards understanding of signal sequence recognition processes during transport.

We observed in both MC4100 and DADE cells expressing intein fused N-terminally with TorA/OmpA signal GFP, cleaved intein is retained in the cytoplasmic fraction and free GFP in the periplasmic fraction (figure 2 in Chapter 2 and figure 2 in Chapter 5). But it is not clear when during the transport the intein gets cleaved and the signal sequences become available for recognition for transport process to continue. Interestingly, we also observed that even as we replaced the intein in the N-terminus of the TorA signal GFP with a mutated intein lacking ability to self-cleave, we again found cleaved intein in the cytoplasmic fraction and free GFP in the periplasmic fraction (data not shown here). This data comes very close to claim the possibility that even though not at the very N-terminus of the construct, the TorA signal sequence would be recognized by the Tat system for the transport of its passenger protein and later cleaved by Signal peptidase I. If the postulate of post-transport cleavage of intein-TorA-GFP is to be proved then this would challenge our current understanding of signal sequence recognition in Tat transport processes. To bridge the gap in our evidence here, the most logical step would be to localize and sequence the expected cleavage product, intein with C-terminal extension of TorA signal sequence clippings resulting from the activities of signal peptidase I in the membrane or signal peptide proteases in the cytoplasm. The intein we use in our constructs can be easily localized and purified from any cellular fractions because of the Chitin binding domain that is engineered in it. The fraction of intein with rather short signal sequence extensions (1-3 kDa) may not be distinguishable from free intein upon gel electrophoretic separation and we may therefore need to employ sensitive detection techniques such as mass spectrometry. This study may also shed light on the fate of Tat signal sequences following the transport process.

I believe that artificial transport substrates like the ones we reported here will aid towards the understanding of transport processes in the coming years.

Curriculum vitae

Anitha Shanmugham was born on March 8th 1980 in Coimbatore, India. Following graduation from Alvernia Matriculation Higher Secondary School in Coimbatore, she joined a 5 year Integrated Master's program in Life Sciences at School of Life Sciences, Bharathidasan University, India. From there, she graduated in May 2002 with Specialization in Biotechnology. During this period she also carried out a 9 months student research project on Caffeine degradation pathways in bacteria at Indian Institute of Sciences, Bangalore under the supervision of Prof. Dr. Ram Rajasekharan, Department of Biochemistry. In June 2002, she started her PhD project in VU-University, Amsterdam under the supervision of Prof. Dr. Holger Lill. Her project aimed to characterize events leading to protein folding and localization in bacterial system. She mainly focused on then recently discovered Twin arginine transport (Tat) pathway which transports folded and even oligomeric proteins across the plasma membrane. Biochemical and biophysical tools were used in *Esherichia coli* to dissect and investigate individual processes step by step. Later in 2007, she joined as a Postdoctoral fellow at Netherlands Kanker Institute, Amsterdam in the Chemical Biology team headed by Dr. Huib Ovaa. There she was involved in development and validation of non-hydrolyzable ubiquitin isopeptide isosteres as probes for studying selectivity in (De)ubiquitination processes. During this period, she had also set up biochemical assay for monitoring the activity of catalytic domain of HAUSP/USP7, a Deubiquitinating enzyme that is involved in stabilization of the tumour suppressor p53. This HAUSP activity assay was proved to be useful tool for high throughput screening of activity inhibitors. She is currently working as a Postdoctoral fellow in the research group of Dr. Iwan de Esch, Department of Medicinal Chemistry, VU-University, Amsterdam where she is involved in setting up biochemical and biophysical (SPR-based) platforms towards screening and development of selective inhibitors for parasitic Phosphodiesterases**.**

Publications

1. **Membrane binding of twin arginine preproteins as an early step in translocation.**

Shanmugham A, Wong Fong Sang HW, Bollen YJ, Lill H. *Biochemistry. 2006 Feb 21;45(7):2243-9.*

2. **DUBs and disease: activity assays for inhibitor development.**

Shanmugham A, Ovaa H. *Curr Opin Drug Discov Devel. 2008 Sep;11(5):688-96. Review.*

3. **Nonhydrolyzable ubiquitin-isopeptide isosteres as deubiquitinating enzyme probes.**

Shanmugham A, Fish A, Luna-Vargas MP, Faesen AC, El Oualid F, Sixma TK, Ovaa H. *J Am Chem Soc. 2010 Jul 7;132(26):8834-5.*

4. **The auto-activation mechanism of USP7/HAUSP by its C-terminal ubiquitin-like domain (HUBL)**

Faesen AC, Dirac AMG, Shanmugham A, Ovaa H, Perrakis A, Sixma TK. *(manuscript submitted to Cell).*

5. **Characherization of the specific interaction of Tat-pathway related chaperones with cognate twin arginine preproteins.**

Shanmugham A, Bakayan A, Völler P, Grosveld J, Lill H, Bollen YJ. *(manuscript submitted to Biochemistry).*

Samenvatting

Het onderzoeken van de bacteriele Tweeling Arginine Transport cascade

Het bestaan van de Tat translocase werd ongeveer 10 jaar geleden ontdekt. Sindsdien is zijn betrokkenheid in verschillende cellulaire processen, zoals energiestofwisseling, celdeling, motiliteit, bacteriële pathogenese en symbiotische stikstoffixatie aangetoont. De Tat translocase verplaatst gevouwen eiwitten over een membraan , zonder de permeabiliteit van het membraan voor kleine moleculen en ionen te verstoren. Hoewel er al enige kennis over de structuur en mechanistische kenmerken beschikbaar is, kunnen er tal van alternatieve modellen worden geconstrueerd op basis van diverse studies op dit gebied en veel vragen blijven onbeantwoord. Vandaag de dag zijn de substraten van de Tat translocase geïdentificeerd en de meest opvallende kenmerken van hun signalering, die transport als doel heeft, zijn bekend. De interacties van de Tat substraat eiwitten, die leiden tot hun export, worden wereldwijd onderzocht.

Het doel van mijn project is het karakteriseren van verscheidene evenementen die leiden tot het transport van Tat, welke ik kan klassificeren als substraat-afhankelijk of translocon-afhankelijk. De aanpak was hoofdzakelijk het uivoeren van *in vitro* experimenten, gezien wij van mening waren dat het efficient is om afzonderlijke evenementen te ontleden in de zeer dynamische Tat signaal cascade en ze één voor één te bestuderen. De in dit proefschrift beschreven experimenten hadden hoofdzakelijk als doel het systematisch begrijpen van biofysische interacties van substraateiwitten met verschillende componenten uit de Tat signaal cascade met gebruik van *Escherichia coli* als model organisme.

In het tweede hoofstuk “Membraanbinding van Tweeling Arginine Pre-eiwitten als een Vroege Stap in Translocatie”, wordt membraanbinding van kunstmatige Tat substraten in real time onderzocht met behulp van Surface Plasmon Resonance Spectroscopy (SPRS). Het is gebleken dat vaste hechting van substraten aan een membraan door signaalpeptiden wordt bewerkstelligd. Wij stellen voor dat dit mogelijk een eerste stap op weg naar transport belichaamt. Voor de volgende stap werden specifieke interacties tussen Tat substraten en hun respectievelijke chaperones getest. Dit werk wordt gepresenteerd in het derde hoofdstuk, “Interacties van Tat-signaal cascade gerelateerde chaperones”. Ik ben van mening dat kunstmatige transport substraten die we in deze twee hoofdstukken beschrijven de komende jaren zullen helpen bij het ontwikkelen van inzichten in transportprocessen.

Om de translocase-georiënteerde evenementen te bestuderen, hebben we geprobeerd om een stabiel Tat translocase complex te ontwikkelen die ons zou helpen om de *in vitro* experimenten te verbeteren. Om dit te bereiken zijn Tat genen die verschillende Tat subdelen van het genoom van het hittebestendige organisme *Aquifex aeolicus* geselecteerd en succesvol gekloneerd in een geschikt membraan eiwit expressie systeem. De expressie en de purificatie van de hittebestendige Tat subdelen zijn bestudeerd. Verder hebben we laten zien dat de

hittebestendige Tat subdelen in staat zijn om de functie van Tat sbudelen in *E coli* over te nemen, zoals beschreven in hoofdstuk vier.

In het vijfde hoofdstuk laten we zien dat het afschermen van Sec-specifieke signaal sequenties ervoor zorgt dat de respectievelijke substraat eiwitten kunnen vouwen en daardoor niet meer herkenbaar zijn voor de Seq cascade. Tot onze grote verrassing bleek dat het transport van de gevouwen Seq substraten wordt herkent door e Tat cascade. Gebaseerd op onze data, maken we de hypothese dat de selectie van transport substraten voor verschillende cascades op een laat post-translatie moment kan plaatsvinden.

Acknowledgement

It is a pleasure to thank all those who made this thesis possible! Foremost, I am heartily thankful to my supervisor, Prof. Holger Lill, for the opportunities and the continuous support he provided for my PhD study from the initial to the final level. Holger, above all and the most needed, you inspired me with your clarity of thoughts and comments with a great sense of optimism in science and as well as in life.

My co-supervisor, Dr. Yves Bollen has made available his support in a number of ways both in research and thesis writing. I owe my deepest gratitude for his patience, motivation, enthusiasm, and immense knowledge. Yves, your ideals and concepts have been of remarkable influence in my career.

I am very grateful to the Prof. Ralf Bernd Klösgen (Martin Luther University), Prof. Antoinette Killian (Utrecht University), Prof. Arnold J M Driessen (University of Groningen), Prof. Wilbert Bitter (Vrije University, Medical Centre) and Dr. Tanneke den Blaauwen (University of Amsterdam) for having participated in the thesis reading committee and for having approved this thesis and for their willingness to extend their participation in the doctoral examination committee. I am equally thankful to Dr. Joen Lurink (Vrije University) for having agreed to participate in the examination committee.

I consider myself fortunate in having colleagues (from Section of Structural Biology, VU) who warmly supported me far beyond call of duty in many different ways apart from what I have recalled here below. Petra Völler, I specially thank for your enormous support in the lab and also for the good time on many occasions outside the lab! Harro Wong Fong Sang, I am very much grateful for helping me with liposome synthesis set up in the lab and more so for offering friendly support during hard times, especially during my house searching period. Dr. Dirk Bald, although we did not work in related projects you have been a breath of fresh air in the department with lively discussions and I am thankful for your timely help and encouragement on various times. Mareike Scholts, I am especially thankful for the friendly efforts you put in to teach me some Dutch during my first year of PhD. Dr. Reiky van Walraven, I will forever remember your immense help on my first day arriving here in Amsterdam, from arranging bank account to finding my accommodation and more. Annelies Jansen, I am thankful for warmly helping me with translations on many occasions. I thank all other group members and students for all their contributions in one way or the other in having made my PhD years comfortable and pleasant one.

During this work I have collaborated with many colleagues for whom I have great regard. I wish to extend my warmest thanks to Dr. Wilfred Röling (Molecular Cell Physiology, VU) for helping me with anaerobic cell culture and Dr. Nathan Brady Röling (Molecular Cell Physiology, VU), for helping with confocal microscopy experiments. I am also grateful to Dr. Connie Jimenez (Molecular and Cellular Neurobiology, VU) for helping with the MALDI-TOF analysis of my protein samples throughout my PhD period.

I am indebted to my student colleagues Rashmi Srivatsava and Adil Bakayan whose contributions were useful to write parts of this thesis.

My special thanks to my (current) remarkable colleagues- Sabrina de Munnik, Petra de Kruijf , Azra Delic and Anne Watts at Pharmacochemie, VU for helping me with the dutch translation of the summary of this thesis.

The external support and encouragement of many friends has been indispensable, and I would like particularly to acknowledge the contribution of Bithiah Grace, Neelayathatchi Subramanian, Patricia Lowery, Sivakumar Solaiappan and Christopher Lowery.

Lastly, and most importantly, I wish to express my love and gratitude to my family. My mother, Nagarathinam S and father, Shanmugham K worked ever industriously to support the family and to provide the best possible environment for me and my brother Thiagu S to grow up and attend school. I could not have asked more of them. Warmest thanks to my husband, Senthilkumar Palanisamy for always being by my side and my son, Rohan Senthilkumar for being my light and joy.